I. PROBLEM STATEMENT

We studied hearing loss algorithms in this project. As the conductive hearing loss is due to sound conducting problem in the ear canal, it can be treated with a linear sound amplifier. However, the other kind of hearing loss, sensorineural hearing loss, which is resulted from the disorder of the middle- or inner-ear, can not be solved merely by a linear amplifier. Sensorineural hearing loss is what we will focus on in this project; and for convenience we mean sensorineural hearing loss simply by hearing loss or sensorineural impairment in later discussion.

The symptoms of sensorineural hearing loss are relatively more complicated. First, the high frequency signal power is diminished in the sound signal perceived by people with sensorineural hearing loss. This makes some phonemes inaudible or can not be heard in a correct way, as the second formant or the higher order formants can not be detected effectively. This problem can be alleviated by frequency compensation in which high frequency component are boosted by different gains according to hearing impairment level (see, e.g., [1][2]). Another symptom of sensorineural hearing loss is that the dynamic range between the weakest sound that can be heard and the most intense sound that can be tolerated becomes much narrower. The weakest sound that can be heard by patient with hearing loss will have a much higher threshold. On the other hand, however, the most intense sound that can be tolerated does not rise as much as threshold of the weakest sound. This narrowed dynamic range will make the normal sound can not be heard to be continuous and smooth by people with sensorineural impairment. A linear amplifier will always make some sound level out of the range, while the non-linear amplification, namely dynamic range compression, can let pass all sound levels into the perception of patients with hearing loss (see, e.g., [1] [3]).

So far we have discussed the two most important aspects about hearing loss, one in frequency domain, and the other in the time domain. Detailed methods to solve these two problems are discussed in Section II and Section III respectively. Additional symptoms of hearing loss include poor frequency resolution and poor time resolution. However, they are beyond the scope of this project for the complexity involved.

II. FREQUENCY COMPENSATION

Patients with hearing loss can not hear some sound, or can not hear the sound in a correct way. This is because the hearing impairment is not of the same level at different frequency band. Hearing impairment at frequency above 700 Hz is much more severe than that in low frequency band. To restore the sound
into a normal level, we need to apply different gain to different frequency band, according to the hearing impairment level there.

This is implemented using a filter-bank approach. For each frame of the speech signal, we first calculate the $N$-point FFT coefficients. Then the signal is split into five 1/3 octave bands (five channels) using triangle windows as depicted in Figure 2, which is very similar to what we do in MFCC. The center frequencies are 250 Hz, 500 Hz, 1 KHz, 2 KHz, and 4 KHz respectively. We use the triangle window rather than the rectangular window to ensure the frequency magnitude of the signal is still continuous after the frequency compensation. In each band, we will amplify the coefficients by a certain gain value that is pre-specified – these gain values should be evaluated according to each patient. We then do an inverse-FFT to the modified coefficients to get the signal in time domain. Figure 1 shows the frequency magnitude of the signal before and after the frequency compensation. We notice that the high frequency bands have much larger gains than that of low frequency bands. From the demo sound produced, we can also hear that the frequency compensated speech have more power distributed in the high frequency.

After that, we employ dynamic range compression to transform the sound level into a range that can be effectively perceived by the patients. This will be discussed in the next section.

**III. Dynamic Range Compression**

Patient with hearing impairment commonly has much higher threshold to hear the weakest sound, but the level of the most intense sound that can be tolerated does not rise by the same amount as the threshold of the weakest sound. With this narrowed dynamic range, the patient can not hear the input sound that is too weak, and may perceived a discontinuous and volatile sound signal. This problem can not be addressed merely by employing a linear amplifier, where no matter how much the sound is amplified there is always some sound level that lies out of the range as in Figure 3. To let the patient hear comfortable and smooth sound, we need to compress the input sound level into the patient’s dynamic range. This dynamic range compression can be viewed as an input-output mapping strategy (I-O curve) shown in Figure 4 (see, e.g., [1][3]). Notice that the beautiful curve in Figure 4 is realized by one of our very talented members using the function which we believe is much better than the linear wide range compression:

$$\frac{1}{1 + e^{-\alpha \tan(d(X-\beta))}},$$

where $\alpha$, $d$, and $\beta$ are user parameters that need to be tuned. As patients have different hearing impairment level, the I-O curve should also be evaluated for each patient. From the I-O curve, we can calculate the gain values for different input SPL (Figure 5).

When implementing the dynamic range compression, the input sound level should be calculated according the past several frames rather than just current one. Otherwise, the noise between voiced
phonemes will be amplified, and the envelopes of the phonemes will be distorted. When determining the gain for the current frame, we calculated the input sound level $\hat{P}(0)$ by averaging the power of the past $N$ frames (including current one) $\{P(n)\}_{n=0}^{N-1}$ weighted by exponential coefficients as

$$\hat{P}(0) = \frac{\sum_{n=0}^{N-1} e^{an} P(n)}{\sum_{n=0}^{N-1} e^{an}}$$  \hspace{1cm} (2)$$

where $a = -0.1$. The gain value for current frame is determined by taking $\hat{P}(n)$ as the input SPL and looking up the gain versus input SPL curves as in Figure 5.

Because the speech signal is processed frame by frame, the signal may not be continuous at the boundaries of the frames. A Chebyshev Type II lowpass filter with cutoff frequency of 8 KHz and -30 dB stopband ripple is used to smooth the output signal and filter out the high frequency noise.

In Figure 6, we show the envelopes for the original sound, frequency compensated sound, dynamic range compressed sound, and the gain values applied by the compressor. In the original sound sound, the signal power from 20 s to 35 s is 12 dB higher than the rest. After the dynamic range compression, the intense sound in the middle is suppressed, and rest weak sound is amplified. In this way, the resulted speech sounds more smooth and more comfortable. We also notice that there are some noises due to the overshoot in the final sound. This is because the gain value for current frame is determined the weighted average of the past 10 frames. If we can look forward a few frames, this overshoot problem should be attenuated, however, it will result in delays for the patient which we should try to avoid.

IV. CONCLUSION

When doing this project, we always wish that if only we know more about the hearing loss and hearing aid facts, as there are so many details that require the information or feedback from the patient. We tried our best to make this project realizable in the real world, if there is any chance. We sincerely hope that there is something in this project that is useful to improve the hearing aid performance.

V. ACKNOWLEDGEMENT

We thank Dr. John Harris and Qing Yang for their directions in our project.

REFERENCES

Fig. 1. Estimated power spectra for a speech frame before and after frequency compensation.

Fig. 2. Frequency bands for filter-bank approach. The center frequencies are 250 Hz, 500 Hz, 1 KHz, 2 KHz, and 4 KHz respectively.
Fig. 3. Loudness dynamic range of ordinary people and patients with hearing impairment. In each of the three parts, the left box indicate the sound levels in a full range; the right box indicate the dynamic range that can be perceived by patients. Without linear amplifier, the patients can not hear the weak sound, while with the linear amplifier, the patients can not hear the intense sound.

Fig. 4. Dynamic range compression. Green curve: high level compression; yellow curve: low level compression; and black curve: wide dynamic range compression.
Fig. 5. Gain versus input SPL that corresponds to I-O curve in Figure 4.

Fig. 6. Waveform profiles of the original sound, frequency compensated sound, and dynamic compressed sound.